INTEREST-FINDER FORM ASSEMBLY: THE INFLUENCE OF PATTERN MATCHING TO THE SELF-DIRECTED SEARCH

Harley Baker and Janet Wall
Defense Manpower Data Center
Seaside, CA

April 1996

DMDC Technical Report 96-001
Approved for public release; distribution is unlimited

The views, opinions, and/or findings contained in this document are those of the authors and should not be construed as the official Department of Defense position unless so designated by other official documentation.

Personnel Testing Division
Defense Manpower Data Center
DoD Center--Monterey Bay
400 Gigling Road
Seaside, CA 93955-6771

INTEREST-FINDER FORM ASSEMBLY: THE INFLUENCE OF PATTERN MATCHING TO THE SELF-DIRECTED SEARCH

Harley Baker and Janet Wall
Defense Manpower Data Center

APRIL 1996

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (*Leave Blank*)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	April 1996	Final

4. TITLE AND SUBTITLE
Interest-Finder Form Assembly: The Influence of Pattern Matching to the Self-Directed Search

5. FUNDING NUMBERS

6. AUTHOR(S)
Harley Baker, Janet Wall

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Personnel Testing Division
Defense Manpower Data Center
DoD Center, Monterey Bay
400 Gigling Road
Seaside, CA 93955-6771

8. PERFORMING ORGANIZATION REPORT NUMBER
DMDC TR 96-001

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Deputy Under Secretary of Defense for Readiness
Office of the Secretary of Defense
Washington, DC 20301

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (*Maximum 200 words*)

Various analyses were performed to compare the results of two form assembly techniques in the development of the Interest-Finder. The technique used to create the actual Interest-Finder selected items so that the interest patterns on the Interest-Finder matched those of the Self-Directed Search. A comparison of items, scales, and forms was made between the actual Interest-Finder and a hypothetical instrument that was produced by eliminating the pattern-matching constraint. Results of the analyses showed that the original form assembly technique used by DMDC test developers produced a superior instrument. The hypothetical instrument exhibited some unacceptable gender differences and did not represent the original test taxonomy nearly as well as the actual Interest-Finder.

14. SUBJECT TERMS

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	

Acknowledgments

Many people contributed to the research that culminated in this paper. We especially wish to acknowledge the contributions made by several members of the Personnel Testing Division of the Defense Manpower Data Center (DMDC). Dr. Lauress Wise, now at HumRRO, wrote the software program that made the initial item selection for the Interest-Finder and provided thoughtful expertise throughout the development of the Interest-Finder. Dr. Iosif Krass wrote, documented, and implemented the new software program that selected the items for inclusion in the Hypothetical Interest-Finder (see Appendix A). Without this program the research reported in this paper would not have been possible. Dr. John Welsh provided comments about some of the statistical analyses. Ms. Gretchen Glick not only meticulously edited various versions of this manuscript as it evolved, but also provided important comments about the content of the paper.

INTEREST FINDER FORM ASSEMBLY: THE INFLUENCE OF PATTERN MATCHING TO THE SELF-DIRECTED SEARCH

EXECUTIVE SUMMARY

The purpose of this technical report is to document the comparisons of two methods of form assembly for the Interest-Finder©, an interest inventory produced by the Defense Manpower Data Center and used as part of the ASVAB Career Exploration Program. Method one, resulting in the operational Interest-Finder, used an item selection procedure that mirrored, to the extent possible, the means and standard deviations of the Self-Directed Search. Method two, resulting in the Hypothetical Interest-Finder, used the same procedures except the pattern matching to the Self-Directed Search was eliminated as a constraint.

A comparison of characteristics of the forms developed using each of these methods was compared by scale or construct and by the entire form. The comparison makes it clear that the two Interest-Finder forms shared a number of common items, though the percentage of overlap was statistically significant for only the Investigative and Artistic areas. It is equally clear that the Interest-Finder tended to provide a more complete coverage of the RIASEC domains than did the Hypothetical Interest-Finder. This was true for the Realistic, Artistic, and Conventional domains. For the other three scales, neither version provided more complete coverage. Even so, this means that the Interest-Finder provided superior content coverage for half of the areas assessed by the measure. While the coefficient alphas for the Hypothetical Interest-Finder scales were uniformly larger than the coefficient alphas for the Interest-Finder scales, these differences were quite negligible--never exceeding .04 in magnitude. Since all of the scales had alphas of at least .91, there is no clear reason to prefer one version over the other. On the other hand, there is a clear reason to prefer Interest-Finder over Hypothetical Interest-Finder when the criterion is that of subgroup differences. In those RIASEC scales for which there were clear differences in the magnitude of the subgroup mean differences between the two versions, Interest-Finder manifested the smaller differences. This was especially true in the comparisons involving the Female sample with the Overall sample. With these considerations in mind, it is reasonable to conclude that the scale-level comparisons endorse Interest-Finder as the better of the two versions.

Conclusions Based on Form Comparisons

Overall, the two versions shared 141 (58.5%) common items, a number that was half-again-larger than what would be expected by chance alone (95 items). Even so, kappa was relatively low ($k = .318$), which suggested that while the overlap was larger than what might be expected due to random selection of the items for the two versions, the two versions did not share common

items significantly beyond what would be expected based on random item selection. From this it was surmised that the two Interest-Finder versions were, indeed, fundamentally different instruments. In this regard it was reasonably safe to conclude that the SDS pattern matching did have a substantial influence on the items that were selected for inclusion in the Interest-Finder.

More importantly, however, was whether such pattern matching adversely affected the coverage of the relevant content areas in the RIASEC constructs. A statistical comparison of the two Interest-Finder forms was not feasible, due to the violation of the minimum expected number of observations per cell. However, an examination of the results of the scale-by-scale comparisons did allow conclusions about the influence of the SDS pattern matching on content considerations. There were no differences in the content coverage on five of the six RIASEC scales. The one difference was on the Realistic scale, for which pattern matching seemed to lead to a more thorough coverage of the content. Based on content considerations, it was reasonable to conclude that pattern matching had a positive, rather than negative, influence on the item-selection process utilized to select items for inclusion in the Interest-Finder.

An analysis of the expected means, standard deviations, and coefficient alphas which were calculated from the item-tryout data for both forms of the Interest-Finder showed that, again, pattern matching either negligibly influenced the results or positively influenced the results.

Together, these findings argued quite substantially that SDS pattern matching led to the creation of a better version of the Interest-Finder than would have been created if pattern matching had not been effected. Based on these considerations, it appeared that the influence of the SDS pattern matching on the Interest-Finder was not negative and that, at worst, the current version of the Interest-Finder suffers no ill effects because of that pattern matching.

INTEREST FINDER FORM ASSEMBLY:
THE INFLUENCE OF PATTERN MATCHING
TO THE SELF-DIRECTED SEARCH

CONTENTS

LIST OF TABLES

INTEREST-FINDER FORM ASSEMBLY: THE INFLUENCE OF PATTERN MATCHING TO THE SELF-DIRECTED SEARCH

The Department of Defense ASVAB Career Exploration Program annually serves hundreds of thousands of high school and postsecondary students in about 14,000 schools nationwide (Wall, 1994). At no cost either to students or participating schools, students complete a battery of measures that assess their aptitudes, interests, and personal preferences and then link their results to characteristics of occupations in both the civilian and military worlds-of-work. These measures include a multiple aptitude battery[1], an interest inventory, and a work values exercise.

In July 1995, the Interest-Finder became the fully operational interest inventory of the ASVAB Career Exploration Program, replacing the Self-Directed Search (SDS; Holland, 1985a). Based on Holland's (1973, 1985b) vocational personality theory--that both people and work environments can be classified according to how well they fit six vocational personality types--the Interest-Finder guides students toward identifying their own vocational personalities, whether they may be Realistic, Investigative, Artistic, Social, Enterprising, and/or Conventional (often referred to as RIASEC types). Three top Interest-Finder RIASEC codes are obtained by each student to use with other personal results (vocational aptitude test battery results and clarification of work values) to identify occupations and careers that are most likely to provide a good match to that student's interests, aptitudes, and values.

The Interest-Finder was developed in three broad phases over a four-year period (Wall, Wise, & Baker, in press). In the first phase, Taxonomy Development and Item Writing, the goals were to specify fully what was to be measured by the Interest-Finder, determine the best type of items to use, write the items, and review the items to ensure their quality. Then, in the Item Tryout and Form Development phase, the items were administered along with the SDS to two large national samples in order to determine the items' psychometric and statistical characteristics and develop a maximum-length form of the Interest-Finder for validation. The third phase, Form Tryout and Validation, focused on further analyzing the psychometric and statistical characteristics of the Interest-Finder items and the scales created from the items. This also included validating the items and the scales against the Strong Interest Inventory (SII; Hansen & Campbell, 1985) and conducting additional validity analyses.

[1] The ASVAB Career Exploration Program includes the Armed Services Vocational Aptitude Battery (ASVAB; U. S. Department of Defense, 1994).

Purpose

The Problem

In the Item Tryout and Form Development phase, an item-selection algorithm was constructed that would determine which items were included in the Interest-Finder. One criterion for the algorithm was that the final Interest-Finder scales would have means and standard deviations proportional to the means and standard deviations of the SDS scales. This would to create an instrument that would reflect a pattern of means and standard deviations similar to an established inventory.

Recently, the Defense Advisory Committee on Military Personnel Testing (DACMPT) expressed concern that the effect of this pattern matching may have been negative; that is, it may have hampered the creation of an optimal measure. This concern gave rise to a specific question: "When developing the final form of the Interest-Finder, what was the effect of pattern matching the Interest-Finder means and standard deviations to those of the Self-Directed Search?"

While a definitive answer to this question is not possible, it is possible to provide a reasonable answer by (a) determining how pattern matching may have affected the Interest-Finder scale means, standard deviations, and coefficient alphas; (b) assessing the means, standard deviations, and coefficient alphas for various subgroups in order to determine how pattern matching may have contributed to any subgroup differences found on the Interest-Finder scales; and (c) determining how pattern matching affected the content of the items selected for inclusion in the Interest-Finder. In addressing the issue raised by the DACMPT, this report will:

- explore the suitable data sources necessary to answer the question;
- explore the potential for form comparison using the item-tryout data;
- review the process used to identify the initial set of Interest-Finder items;
- describe salient pattern-matching characteristics of the software program that selected the items;
- describe the characteristics of a new software program that does not implement pattern matching;
- describe the criteria used to compare the two forms of the Interest-Finder;
- compare the Interest-Finder scales created by the two software programs;
- compare the Interest-Finder forms created by the two software programs; and
- provide a summary comparison between the Interest-Finder forms.

For ease of explanation, the current form of the Interest-Finder is referred to as the Interest-Finder, and the set of items selected for inclusion by the software program that eliminated the pattern matching is referred to as the Hypothetical Interest-Finder. While the items that comprise the Interest-Finder subsequently were placed together into a form for a final validation study, the items selected for inclusion into Hypothetical Interest-Finder have not been placed into a single form and administered to respondents.

Appropriate Data Sources

The data collected in the second (Item Tryout and Form Development) and third (Form Tryout and Validation) phase of the development of the Interest-Finder can be used to provide (a) a basis with which to explore the issue and (b) an answer to the question posed by the DACMPT. From the Item Tryout and Form Development phase, the data consist of the item-level means and standard deviations and the item-to-SDS scale correlations. With these data, it is possible to construct estimates of what the means, standard deviations, and coefficient alphas of the resultant scales would be (Lord & Novick, 1968). From the Form Tryout and Validation phase, the data consist of the item-level means and standard deviations and the item-to-Interest-Finder scale correlations. As can be seen in Table 1, which identifies the data to be used in conducting the comparisons, there are no scale-level data for the Hypothetical Interest-Finder; therefore, the items "selected" for this measure do not really constitute a set of scales or a form.

Table 1
Data Sources Necessary for a Comparison Between the Interest-Finder and the Hypothetical Interest-Finder

	Interest-Finder		**Hypothetical Interest-Finder**	
Available Data Source	**Available Item-Level Data**	**Available Scale-Level Data**	**Available Item-Level Data**	**Available Scale-Level Data**
Item-Tryout Data	Mean	Estimated Mean	Mean	Estimated Mean
	Standard Deviation	Estimated Standard Deviation	Standard Deviation	Estimated Standard Deviation
	Item-to-SDS Scale Correlations	Estimated Coefficient Alpha	Item-to-SDS Scale Correlations	Estimated Coefficient Alpha
Form-Tryout Data	Mean	Actual Mean		
	Standard Deviation	Actual Standard Deviation		
	Item-to-Interest-Finder Scale Correlations	Actual Alpha		

The item-tryout data were obtained from a sample of students (N = 3,233) from 61 schools across 24 states, the District of Columbia, and Puerto Rico. The sample consisted of approximately equal numbers of males (48%; n = 1,548) and females (52%; n = 1,685). Ethnically, this sample was fairly diverse: Caucasian students comprised 56% of the sample, Blacks comprised 27%, Hispanic Americans comprised 7%, Asian Americans comprised 5%, and students from other racial/ethnic descents comprised 3% of the sample. Based on the National Center for Educational Statistics formulation (Eagle, 1989), the sample exhibited a wide

range in socioeconomic status (SES; -3.52 to 3.12) and a mean indicative of a middle-class sample (M = .22, SD = .90). A value of 0 indicates "middle class" in this scale, with positive values indicating higher-than-average SES and negative values indicating lower-than-average SES.

Because it was not possible to administer all 607 of the tryout items to each student, the new items were divided into six tryout forms. Each form contained (a) a series of background items from which geographic diversity and SES could be determined, (b) the SDS, and (c) about 100 Interest-Finder tryout items. Data were collected in schools from April through June of 1993. The number of students who completed each form ranged from 525 to 556, with a median sample size of 537 respondents per form.

In comparison, the Form Tryout and Validation sample consisted of 1,319 high school students from 22 schools across 20 states. The number of participants per school ranged from 13 to 126, with an average of 60 participants per school. Males (45%; n = 591) and females (55%; n = 722) were fairly equally represented in this sample. Students' ages ranged from 13 to 19, with an average age of 16.4 years (SD = 1.1 years). Most students were sophomores (46%), juniors (24%), or seniors (28%), with only a few freshmen (2%) in the sample. While a large proportion of the sample was Caucasian (63%), there were Blacks (9%), Hispanics (13%), Asian Americans or Pacific Islanders (6%), and Native Americans (5%) in the sample. The sample exhibited a wide range in SES (-2.13 to 2.92) and an average indicative of a middle-class sample (M = .26, SD = .87).

It appears that the available data lend themselves to two types of comparisons. The first is a comparison between estimated and observed scale characteristics. It is possible to compare the estimates of the Interest-Finder scale characteristics (estimated from the item-tryout data) with the observed scale characteristics (calculated from the form-tryout data). If these comparisons indicate that the estimated scale characteristics are similar to the observed scale characteristics, the second type of comparison can be conducted. In the second type of comparison, the estimated scale characteristics of the Interest-Finder are compared with the estimated scale characteristics of the Hypothetical Interest-Finder. Just as the item-tryout data were used to construct estimates of the Interest-Finder scale characteristics (means, standard deviations, and coefficient alphas), the item-tryout data could also be used to construct estimates of the scale characteristics of the Hypothetical Interest-Finder. A comparison of these estimated scale characteristics could then be conducted to provide valuable information about the effects of pattern matching on the Interest-Finder.

Estimated and Observed Interest-Finder Scale Characteristics

The first type of comparison concerns the estimated and observed Interest-Finder scale characteristics. Table 2 reports the estimated and observed scale means, standard deviations, and coefficient alphas. As can be seen, there were some large differences between the estimated Interest-Finder scale characteristics (as estimated from the item-tryout data) and the observed scale characteristics (as calculated from the form-tryout data).

Table 2

Estimated Scale Characteristics vs. Observed Scale Characteristics for the Interest-Finder

	Estimated Scale Characteristics			Observed Scale Characteristics			Observed - Estimated Differences			
Interest-Finder Scale	Mean	SD	α	Mean	SD	α	Mean	SD	α	Cohen's *d*
Realistic Scale										
Overall Sample	9.48	8.96	.94	10.16	8.80	.93	.68	−.16	−.01	.08
Female Sample	6.11	6.60	.91	7.51	7.14	.91	1.40	.54	.00	**.20**
Black Sample	9.02	8.54	.93	8.17	7.95	.93	−.85	−.59	−.01	−.11
Hispanic Sample	9.14	8.17	.92	9.44	8.50	.93	.30	.33	.01	.04
Low SES Sample	9.71	8.83	.94	10.72	8.99	.93	1.01	.16	.00	.11
Investigative Scale										
Overall Sample	11.42	8.96	.94	14.60	10.67	.95	3.18	1.71	.01	**.30**
Female Sample	10.45	9.79	.95	13.94	10.28	.94	3.49	.49	.00	**.34**
Black Sample	10.84	9.56	.94	12.48	10.28	.95	1.64	.72	.00	.16
Hispanic Sample	12.38	9.63	.94	15.24	10.68	.95	2.86	1.05	.01	**.27**
Low SES Sample	10.31	9.57	.94	14.00	10.38	.94	3.69	.81	.00	**.36**
Artistic Scale										
Overall Sample	11.06	9.58	.94	12.97	8.98	.94	1.91	−.60	.00	**.21**
Female Sample	11.48	10.06	.95	14.06	9.91	.94	2.58	−.15	−.01	**.26**
Black Sample	11.39	8.56	.92	13.91	9.16	.92	2.52	.60	.00	**.28**
Hispanic Sample	11.92	8.77	.92	12.89	9.84	.94	.97	1.07	.02	.10
Low SES Sample	10.48	9.02	.93	12.30	9.48	.93	1.82	.46	.00	.19
Social Scale										
Overall Sample	13.18	9.59	.93	14.61	10.21	.94	1.43	.62	.01	.14
Female Sample	15.72	9.24	.92	17.47	10.09	.93	1.75	.85	.01	.17
Black Sample	14.17	9.20	.92	15.50	9.94	.93	1.33	.74	.01	.13
Hispanic Sample	13.16	9.88	.94	16.30	10.70	.94	3.14	.82	.00	**.29**
Low SES Sample	12.76	9.39	.93	14.74	10.13	.94	1.98	.74	.01	**.20**
Enterprising Scale										
Overall Sample	12.80	10.68	.95	14.55	11.22	.95	1.75	.54	.00	.16
Female Sample	12.99	10.62	.95	14.34	10.93	.95	1.35	.31	.00	.12
Black Sample	14.35	10.87	.95	15.59	10.66	.95	1.24	−.21	.00	.12
Hispanic Sample	12.60	9.59	.94	15.20	11.18	.95	2.60	1.59	.01	**.23**
Low SES Sample	12.20	10.81	.95	14.04	11.16	.95	1.84	.35	.00	.16
Conventional Scale										
Overall Sample	9.67	10.56	.96	8.97	13.30	.96	−.70	2.74	.00	−.05
Female Sample	10.88	11.25	.97	10.09	10.51	.96	−.79	−.74	−.01	−.08
Black Sample	12.11	11.00	.96	10.07	10.09	.95	−2.04	−.91	−.01	**−.20**
Hispanic Sample	9.84	9.60	.95	10.55	10.77	.96	.71	1.17	.01	.07
Low SES Sample	10.20	10.54	.96	9.94	10.72	.96	−.26	.18	.00	−.02

Note. Estimated scale characteristics based on item-tryout data. Observed scale characteristics based on form-tryout data. The one-sample, modified Cohen's *d*, an effect-size statistic, was based on the item tryout and form tryout means and the form tryout standard deviation. A *d* of .20 or larger suggests an important difference between the form-tryout (observed) and item-tryout (expected) means and is in boldface type.

It is helpful to understand why the estimated scale means and standard deviations are so different from the observed scale means and standard deviations. The basis for these scale-level differences can be seen in the item-level endorsement rates.[2] A comparison of the item-tryout endorsement rates and the form-tryout endorsement rates revealed significant differences on 97 of the 240 Interest-Finder items. When compared with the item-tryout endorsement rates, 12 items exhibited significantly lower form-tryout endorsement rates and 85 items exhibited significantly higher form-tryout endorsement rates. Based on these differences, it would be reasonable to expect that the observed scale means would differ substantially from the estimated scale means. This was the case, as it is apparent that the observed means diverged considerably from the estimated means. While Table 2 reports the arithmetic differences, the absolute differences between the estimated and observed means range from a low of .26 to a high of 3.69, with a median value of 1.70. The differences between the estimated and observed standard deviations also shows a fair amount of divergence, with the absolute differences ranging from .15 to 2.74, with a median of .61.

One way to evaluate the importance of the differences between the observed and the estimated scale statistics is to assess the magnitude of the differences in terms of their effect sizes. A modified form of Cohen's *d* was calculated for each of the mean comparisons reported in Table 2. While there is no standard amount that constitutes a threshold point, Cohen (1988) indicates that effect sizes smaller than .20 generally are not sufficiently large to warrant serious consideration, while effect-size estimates of .20 or more are large enough to warrant consideration. Using this as a criterion, 12 of the 30 mean comparisons had observed means that were sufficiently different from the estimated means that they may warrant some concern.

Undoubtedly, there are several probable reasons for the differences between the estimated and the observed values; one may be fatigue effects associated with the differential placement of the Interest-Finder items in the two studies (Anastasi, 1988; Cronbach, 1990). In the item-tryout study, the items were presented after students had already completed the demographic items and the SDS. In the form-tryout study, the Interest-Finder items were presented after the demographic items but before the SII. This may have caused relatively larger fatigue effects in the item-tryout study than in the form-tryout study. Such fatigue, if present, would probably have tended to lower the scores in the item-tryout study relative to the form-tryout study. Such lowered item-level scores would be reflected in lowered scale scores and standard deviations as well.

Another potential reason is also based on the way in which the Interest-Finder items were presented to the participants. In the item-tryout study, the items were not presented in the context of a completed instrument. Though grouped together by RIASEC area, the items did not have the

[2] It is also possible that the differences might be attributable to the fact that the item-tryout data did not provide all of the item-level data required by the formulas reported by Lord and Novick (1968). While the item-tryout data were able to provide item-level means and standard deviations, they were not able to provide item-to-Interest-Finder scale correlations. Instead, these data provided item-to-SDS scale correlations. These values were then substituted for the item-to-Interest-Finder scale correlations in the formulas. To determine the effect of this substitution, the item-to-Interest-Finder scale correlations from the form-tryout phase were compared with the item-to-SDS scale correlations from the item-tryout phase. A 2 X 6 mixed ANOVA (item correlation: Interest-Finder, SDS X scale: R, I, A, S, E, C) revealed no significant differences either for the untransformed correlations [$F(1, 234) = .58$, $p = .458$] or for the Fisher-transformed correlations [$F(1, 234) = 2.97$, $p = .086$. These ANOVA results support the conclusion that the substitution of the item-to-SDS scale correlations for the item-to-Interest-Finder scale correlations had, at most, a negligible effect on the estimates of the scale characteristics.

look or feel of a completed instrument as they did in the form-tryout study. This may have made the items in the form-tryout study more salient than when presented in the item-tryout study. Such increased salience is consistent with increased levels of motivation, which generally increases test scores (Cronbach, 1990).

A third reason for the difference may lie in the demographic characteristics of the two samples. For example, the form-tryout sample had significantly fewer seniors than did the item-tryout sample [$\chi^2(1, N = 4{,}539) = 34.69, p < .0001$]. From the form-tryout data it has been determined that compared with freshmen, sophomores, and juniors, seniors have significantly higher scores on four of the six Interest-Finder scales (Wall, Wise, & Baker, unpublished data). The lower means from the item-tryout data are consistent with this finding. Similarly, the item-tryout study sample had a significantly larger proportion of respondents between the ages of 13 and 16, while the form-tryout sample had significantly more 17 through 19 year olds [$\chi^2(1, N = 4{,}428) = 21.24, p < .002$], another situation which would tend to produce lower means from the item-tryout data than from the form-tryout data (Wall, Wise, & Baker, in press).

Even though the differences for the means and standard deviations are larger than what one might hope, the pattern of the estimated and observed scale means was quite similar. These similarities were captured in the correlation between the estimated and observed sets of scale means. These ranged in value from .77 to .98, as reported in Table 3. A similar procedure was employed to assess the similarities between the estimated and the observed standard deviations. The patterns here, too, were quite consistent, with the correlations ranging from .72 to .94. These high correlations indicate that while there were relatively large absolute differences between the estimated and observed means and standard deviations, the relationships among the estimated scale values substantially mirrored the relationships among the observed scale values.

Table 3
Pattern Correlations Between the Estimated and Observed Means and Standard Deviations

	Correlation Between Estimated and Observed Scale Means		Correlation Between Estimated and Observed Scale SDs	
Sample	*r*	*p*	*r*	*p*
Overall Sample	.90	.003	.72	.035
Female Sample	.91	.002	.94	.001
Black Sample	.83	.010	.76	.023
Hispanic Sample	.98	.001	.94	.001
Low SES Sample	.77	.022	.94	.001

Note. Probability values (*p*) based on one-tailed tests of significance, with $df = 5$ for each correlation (*r*).

Taken together, the absolute differences presented in Table 2 and the pattern similarities reported in Table 3 lead to three major findings: (a) the observed scale alphas from the form-tryout phase were estimated accurately from the item-tryout data; (b) neither the observed scale means nor standard deviations from the form-tryout phase could be estimated accurately from the item-tryout data; and (c) there was a high degree of pattern similarity between the estimated and observed scale means and standard deviations. Based on these findings, it appeared that the item-tryout data would be useful for estimating coefficient alpha. The use of the estimates of the scale

means and standard deviations in terms of patterns, rather than in terms of absolute quantities, was also justified. Consequently, Interest-Finder form comparisons based on the estimates calculated from the item-tryout data needed to focus on these patterns and relationships, rather than on the absolute magnitude of the estimates themselves.

Method

Interest-Finder Item Selection Procedures

Interest-Finder item selection proceeded in two stages. In the first stage, general item-level screenings were conducted to identify the items that met the statistical requirements for inclusion in the inventory. Three screening criteria were designated, and items that failed to meet these criteria were deleted from further consideration: (a) The item endorsement rate needed to be between 10% and 75%; (b) each item needed to correlate more highly with its target SDS scale than with any other SDS scale; and (c) gender and ethnicity endorsement rate differences needed to be less than 40%.

In the second stage, a software program was written and implemented to select the items that would produce scales that would roughly match the list of specified scale characteristics. The program selected an initial group of items for each scale and then considered all possible replacements by deciding whether each replacement led to a better form. The program continued until it selected the "best" set of items based on the desired scale characteristics. Nine criteria, or goals, were specified for the scales. These criteria differed in their degree of importance in the item-selection algorithm. Each criterion was assigned a weight (w) that determined its relative importance in the item-selection process. Of the nine criteria, five were associated with endorsement rates and four were associated with scale intercorrelations. The five endorsement-rate criteria were designed to create Interest-Finder scales that exhibited the following:

- means and standard deviations proportional to SDS means and standard deviations (w = 1.0);
- minimal gender endorsement-rate differences (w = .5);
- minimal Caucasian-Black endorsement-rate differences (w = .5);
- minimal Caucasian-Hispanic endorsement-rate differences (w = .5); and,
- minimal endorsement-rate differences due to SES (w = .2).

The magnitude of the weights for each of these criteria indicate that one of the most important endorsement-rate criterion was the development of scales with means and standard deviations that would be proportional to the means and standard deviations of the SDS scales. This criterion established that the RIASEC profile of means of the new scales would be similar to the RIASEC profile of means of the SDS scales. Reducing endorsement-rate differences due to gender, ethnicity, and SES were also important criteria. It was hoped that the selection of items that showed smaller gender, ethnicity, and SES differences would result in scales equally appropriate for males and females, as well as for various ethnic and SES groups.

The four correlational criteria were designed to select items that exhibited the following:

- high item-to-SDS scale correlations (w = 1.0);
- minimal item-to-opposite-SDS scale correlations (w = .5);
- minimal item-to-alternate-SDS scale correlations (w = .3); and,
- minimal item-to-adjacent-SDS scale correlations (w = .2).

Based on the magnitude of the weights, the most important correlational criterion was the item-to-target scale correlations. Using Holland's hexagonal model as a reference, the other three criteria were to minimize opposite-scale correlations, alternate-scale correlations, and adjacent-scale correlations. It was hoped that these criteria would result in scales that approximated the hexagonal model proposed by Holland.

The initial set of items that were selected by the software program were then further reviewed and screened to ensure that the items constituted a balanced representation of the content taxonomy. Based on the detailed content taxonomy, these screenings were conducted to ensure that the scales provided appropriate coverage of the RIASEC constructs, an important content validity consideration. As a result of these further reviews and screenings, only 15 items (6.3%) were replaced. The final validation form consisted of items that were conceptually and psychometrically sound and would facilitate a complete assessment of the RIASEC domains. As a result of these screenings, the 240-item Interest-Finder tryout form was finalized.

Pattern Matching and the Original Item-Selection Algorithm

In essence, item selection was achieved by defining a weighted function of the items selected for inclusion in a scale. The formula for this function is as follows:

$$\text{wfunc} = \sum [\text{wt1(i)} * (\text{stmean} - \text{targmn(i)})^2 + \text{wt2(i)} * (\text{stvar} - \text{targsd(i)})^2]$$

where,

wt1(i) = the designated weight for the importance of the ith implemented criterion,
stmean = the scale mean with the item included,
targmn = the appropriate SDS scale mean,
wt2(i) = the designated weight for the importance of the ith implemented criterion,
stvar = the scale standard deviation with the item included, and
targsd = the appropriate SDS scale standard deviation.

The program then calculated this weighted function for each possible scale form that could be created from the items within each of the three item types (activities, training, and occupations). The combination of items that created the scales and minimized this function constituted the final scales. As can be seen, the use of this weighted function made the pattern-matching criterion the basis for the entire program. Unlike a weighting scheme where the weights could be set to zero to remove their influence, the pattern-matching criterion could not be ignored.

Implementation of a New Item Selection Algorithm

The original item-selection program, written in BASIC, was analyzed to determine how it could be modified either to (a) remove the pattern matching from the algorithm, or (b) remove the influence of the pattern matching from the algorithm. After careful analysis, however, it was determined that modifying the original program to remove the pattern-matching criterion was not feasible because of the process it used to select items for inclusion in the scales. Hence, it was decided that a new program needed to be constructed that would conduct Hypothetical Interest-Finder item selection by implementing the same selection criteria as the original program, but without pattern matching to the SDS means and standard deviations (see Appendix A for a detailed description of this program).

A different weighted function was defined that did not depend on SDS pattern matching, but the remaining item-selection criteria were implemented in a way that would select the same items and, thus, mirror the original program. The new program was utilized to select items for a hypothetical version of the Interest-Finder, herein called Hypothetical Interest-Finder. Like the Interest-Finder, the Hypothetical Interest-Finder consists of six scales; each scale assesses one of the six RIASEC domains and is composed of 14 Activities items, 12 Training items, and 14 Occupation items.

Criteria for Scale Comparison

The two forms of the Interest-Finder scales were compared according to four criteria. The first criterion was the degree of item overlap, or number of common items, between the two forms as assessed by calculating kappa. A high value for kappa indicated substantial overlap between the two scales and that such overlap was beyond what would be expected based on chance alone. A high value for kappa would also suggest that the two scales have similar item content. A low value for kappa would suggest that the two scales share few items in common. The statistical significance of the kappa was used to determine if the magnitude of the kappa was high or low.

The second criterion was the degree of content coverage provided by the two Interest-Finder forms. While it may be argued that this is an unfair comparison, since Interest-Finder items were screened and Hypothetical Interest-Finder items were not screened, only 15 of the 240 Interest-Finder items were replaced as a result of the screening. Since this is such a low replacement rate (6.3%), content coverage is a reasonable area of comparison between the two Interest-Finder forms.

The third criterion was more statistical in nature, based on the means, standard deviations, and coefficient alphas calculated from the item-tryout data. While these estimates do not allow for tests of statistical significance,[3] they do provide considerable comparative information.

[3] The scale "statistics" are merely estimates. Such estimates are not amenable to significance testing because such testing presupposes that there is some "population" from which the initial data were sampled, and that it is this population to which we want to generalize the findings. As the scale statistics are estimates derived from data that were not obtained from a single population,there is no real population for whom the estimates are relevant. Hence the use of inferential statistics is not appropriate.

The final criterion, also statistical, focused on the degree of subgroup differences in both forms of the Interest-Finder. Testing these differences for statistical significance is also not appropriate, given the nature of the data.

Results

Realistic Interest-Finder and Hypothetical Interest-Finder Scale Comparison

Cohen's kappa was computed to assess the degree to which the items selected by the pattern-matching algorithm (Interest-Finder) were also selected by the non-pattern-matching algorithm (Hypothetical Interest-Finder). The value of kappa was exceedingly low (k = .025) and strongly suggested that the two Interest-Finder Realistic scales had no more common items than one would expect by chance alone. The items that comprise both scales are reported in Table 4. Of the 40 items selected for these scales, there were only 16 (40%) common items. While this may seem to be a reasonable number, it needs to be interpreted in light of the total number of items available for selection. From this perspective, the two scales selected 16 common items out of a total of 104 items, only about 15%. Random selection of the items for the two scales would have led to 15 common items--one fewer than was found.[4]

Having so few common items raises a number of possible concerns. One concern is that the two scales might be assessing different aspects of the construct(s) that underlie the Realistic dimension. If so, this would suggest that the two scales could be quite dissimilar from each other. Another concern is that the two scales might exhibit different psychometric properties. Fortunately, both of these concerns can be addressed, and in addressing these concerns a conclusion can be drawn as to which is the better scale.

---

[4] The expected number of common items due to random item selection can be calculated from the formula:

$$CI = N * \frac{n1}{t1} * \frac{n2}{t2}$$

where,

CI = expected number of common items,
N = number of scale items (40),
n1 = number of scale items in Interest-Finder (40),
t1 = number of possible items for inclusion in the Interest-Finder scale,
n2 = number of scale items in Hypothetical Interest-Finder (40), and
t2 = number of possible items for inclusion in the Hypothetical Interest-Finder scale.

Because the number of tryout items differed among the RIASEC scale areas, the expected number of common items will differ accordingly among the RIASEC scales.

Table 4

Interest-Finder and Hypothetical Interest-Finder Realistic Scale Items

Item Type	Interest-Finder	Hypothetical Interest-Finder
Act	**Drive a tractor**	**Drive a tractor**
Act	**Operate timber and logging equipment**	**Operate timber and logging equipment**
Act	**Repair broken locks**	**Repair broken locks**
Act	**Repair telephone lines**	**Repair telephone lines**
Act	**Replace broken windows**	**Replace broken windows**
Act	Apply wood stains and varnishes to furniture	Assemble a stereo system
Act	Connect a VCR	Build houses
Act	Drive a fire engine or ambulance	Operate a forklift
Act	Frame a house	Operate heavy construction equipment
Act	Operate a lawn mower	Repair a lawn mower
Act	Refinish the floor in a house	Repair a wooden fence
Act	Replace a watch battery	Replace a broken light switch
Act	Take care of domestic animals	Use a carpenter's level
Act	Use a battery tester	Work with carpentry tools
Train	**Automobile tune-up**	**Automobile tune-up**
Train	**Building a deck for a house**	**Building a deck for a house**
Train	**Installing fire alarms**	**Installing fire alarms**
Train	**Repairing small home appliances**	**Repairing small home appliances**
Train	**Woodworking**	**Woodworking**
Train	Bookshelf construction	Automotive body work
Train	Controlling garden pests	Car engine repair
Train	Furniture repair	Construction of wooden furniture
Train	Installing telephones	Elevator repair
Train	Painting a house (interior or exterior)	Finishing a basement
Train	Planting or harvesting of farm crops	Installation of central air conditioning systems
Train	Restoration of antique furniture	Installing wiring in a house
Occ	**Industrial machinery mechanic**	**Industrial machinery mechanic**
Occ	**Plumber**	**Plumber**
Occ	**House painter**	**House painter**
Occ	**Farm equipment mechanic**	**Farm equipment mechanic**
Occ	**Woodworking machine operator**	**Woodworking machine operator**
Occ	**Television repairer**	**Television repairer**
Occ	Computer repairer	Aircraft electrician
Occ	Emergency vehicle driver	Automobile stereo installer
Occ	Farmer	Automotive body repairer
Occ	Gardener	Carpet installer
Occ	Landscaper	Diesel engine mechanic
Occ	Telephone repairer	Electronic weapons systems repairer
Occ	Tree trimmer	Heavy machine operator
Occ	Wallpaper hanger	Radar equipment repairer

Note. Items common to both scales are in boldface type. Act = Activities item. Train = Training item. Occ = Occupation item.

The major content categories for the Realistic scale are reported in Table 5. When the items that comprise the two scales were categorized according to that taxonomy, a significant difference was found between the coverage provided by each scale [$\chi^2(3, N = 80) = 9.49, p = .023$]. It is fairly apparent that the Hypothetical Interest-Finder Realistic scale, which was created without SDS pattern matching, failed to provide an adequate representation of the taxonomy. Such loss of adequate representation seriously calls into question the adequacy of this scale.

Table 5
Content Comparison of the Two Realistic Scales

	Interest-Finder Items		Hypothetical Interest-Finder Items	
Realistic Content Areas	Number	Percent	Number	Percent
Animals and Agriculture	7	17.5%	0	0.0%
Building and Construction	14	35.0%	12	30.0%
Electricity and Electronics	9	22.5%	10	25.0%
Machinery and Engines	10	25.0%	18	45.0%

The estimated means, standard deviations, and coefficient alphas (Lord & Novick, 1968) for both Realistic scales are reported in Table 6 for the Overall sample and for various important subgroups.

Table 6
Estimated Realistic Scale Means, Standard Deviations, and Coefficient Alphas

	Interest-Finder			Hypothetical Interest-Finder		
Sample Groups	Mean	SD	Alpha	Mean	SD	Alpha
Overall Sample	9.48	8.96	.94	9.39	11.60	.97
Female Sample	6.11	6.60	.91	3.94	6.50	.94
Black Sample	9.02	8.54	.93	8.26	10.65	.97
Hispanic Sample	9.14	8.17	.92	8.64	10.39	.96
Low SES Sample	9.71	8.83	.94	9.54	11.60	.97

There are several conclusions that can be drawn from the sample mean estimates. First, the Interest-Finder means tended to be higher than the Hypothetical Interest-Finder means. Second, both sets of means appeared to have a similar pattern. This is most easily seen by the rank order of the subgroup means for each Interest-Finder form: the rankings were the same for both Interest-Finder Realistic scales. Third, although the rankings were the same, there was a much greater degree of difference between the subgroup means and the Overall sample mean on the Hypothetical Interest-Finder Realistic scale than on the Interest-Finder Realistic scale. This was most clearly seen for the Female sample, where the difference from the Overall sample was quite substantial (3.37 and 5.45 respectively). Among the standard deviations, there were also some consistent differences. With the Female sample exception, the Interest-Finder Realistic scale standard deviations were substantially smaller (between 25-30%) than those for the Hypothetical Interest-Finder Realistic scale. Interest-Finder alphas were all smaller than the Hypothetical Interest-Finder alphas by about .03 or so.

Taken together, the findings suggest that SDS pattern matching may have served to improve the quality of the Interest-Finder Realistic scale. From a taxonomy perspective, it led to a broader coverage of the content, as Hypothetical Interest-Finder had significant gaps in the content coverage. From a quantitative perspective, pattern matching led to a scale that seemed not to exhibit the large subgroup differences found on the Hypothetical Interest-Finder scale. This was most dramatically seen in the comparison of the Female sample mean with the Overall sample mean. For the Interest-Finder scale, the difference was sizable, 3.37 points. However, for the Hypothetical Interest-Finder scale, the difference was over half again larger, 5.45 points. The differences between the Overall sample and the remaining subgroups, although not as large, were all quite substantial, with the larger differences on the Hypothetical Interest-Finder scale.

Based on these comparisons, it appears that the effect of the pattern matching was positive and that the pattern-matched Interest-Finder Realistic scale (Interest-Finder) was superior to the non-pattern-matched scale (Hypothetical Interest-Finder).

Investigative Interest-Finder and Hypothetical Interest-Finder Scale Comparison

The items that comprised both Investigative scales are reported in Table 7. Cohen's kappa was computed to assess the degree to which the items selected for the Interest-Finder Investigative scale were also selected for the Hypothetical Interest-Finder Investigative scale. The value of kappa was moderately large ($k = .524$), suggesting that the two Investigative scales shared significantly more items than would be expected by chance alone. Of the 40 items selected for each scale, there were 28 (70%) common items. Random selection would have led to 15 common items, so it appears that there is a high degree of item-overlap between the two Interest-Finder Investigative scales. Further, the overlap is fairly consistent across the three item types (activities, training, and occupations).

While issues surrounding similarity of content coverage are still important, the relatively high degree of overlap between the two forms of the Interest-Finder Investigative scales suggest that this is probably less of an issue than it was for the Realistic scale. This is especially true since the overlap was consistent across item types. As such, there is also less concern that the two scales might be assessing different aspects of the construct(s) that underlie the Investigative dimension, or that the two scales might exhibit largely different psychometric properties, such as coefficient alphas.

Table 7

Interest-Finder and Hypothetical Interest-Finder Investigative Scale Items

Item Type	Interest-Finder	Hypothetical Interest-Finder
Act	**Conduct research to improve solar power**	**Conduct research to improve solar power**
Act	**Find the area of a triangle**	**Find the area of a triangle**
Act	**Learn about chemical compounds**	**Learn about chemical compounds**
Act	**Learn scientific ways to help protect the environment**	**Learn scientific ways to help protect the environment**
Act	**Prove geometry theorems**	**Prove geometry theorems**
Act	**Study about new sources of energy**	**Study about new sources of energy**
Act	**Study chemical reactions**	**Study chemical reactions**
Act	**Study plants under a microscope**	**Study plants under a microscope**
Act	**Study the effects of radiation on plants**	**Study the effects of radiation on plants**
Act	Examine the ruins of an ancient temple	Explain how satellites work
Act	Operate a computer to solve complex math problems	Identify different types of science lab equipment
Act	Study how diseases are spread	Take a Statistics class
Act	Study marine life	Use a prism to study light
Act	Study the ecosystem of a coral reef	Write scientific reports
Train	**Conducting biology experiments**	**Conducting biology experiments**
Train	**Conducting chemistry experiments**	**Conducting chemistry experiments**
Train	**Conducting lab experiments**	**Conducting lab experiments**
Train	**Conducting physics experiments**	**Conducting physics experiments**
Train	**Geometry**	**Geometry**
Train	**Metric system**	**Metric system**
Train	**Operating a telescope**	**Operating a telescope**
Train	**Scientific methods**	**Scientific methods**
Train	Algebra	Earth sciences
Train	Animal anatomy	Using mathematics to predict economic growth
Train	Animal behavior in the wild	Water pollution
Train	Diseases and their cures	Weights and measures
Occ	**Computer scientist**	**Computer scientist**
Occ	**Earth scientist**	**Earth scientist**
Occ	**Ecologist**	**Ecologist**
Occ	**Environmental research scientist**	**Environmental research scientist**
Occ	**Experimental research worker**	**Experimental research worker**
Occ	**Geographer**	**Geographer**
Occ	**Marine biologist**	**Marine biologist**
Occ	**Medical examiner**	**Medical examiner**
Occ	**Nuclear engineer**	**Nuclear engineer**
Occ	**Science laboratory assistant**	**Science laboratory assistant**
Occ	**Scientist**	**Scientist**
Occ	Animal scientist	Archaeologist
Occ	Eye doctor	Inventor of scientific products
Occ	Veterinarian	Mathematician

Note. Items common to both scales are in boldface type. Act = Activities item. Train = Training item. Occ = Occupation item.

Seven taxonomic areas defined the Investigative dimension (see Table 8). When the items that comprise the two scales were categorized according to the Investigative content taxonomy, as expected, no significant differences were found [$\chi^2(6, N = 80) = 4.72, p = .580$]. While there are some apparent differences in the coverage, especially for Life Science, it appears that the content coverage is approximately the same for each Investigative scale.

Table 8
Content Comparison of the Two Investigative Scales

Investigative Content Areas	Interest-Finder Items		Hypothetical Interest-Finder Items	
	Number	Percent	Number	Percent
Behavioral and Social Science	1	2.5%	1	2.5%
Computer Science	2	5.0%	1	2.5%
Engineering	1	2.5%	1	2.5%
General Science	6	15.0%	9	22.5%
Life Science	16	40.0%	8	20.0%
Mathematics	5	12.5%	8	20.0%
Physical Science	9	27.5%	12	30.0%

The estimated means, standard deviations, and coefficient alphas (Lord & Novick, 1968) for the two Investigative scales are reported in Table 9, both for the Overall sample and for the important subgroups.

Table 9
Estimated Investigative Scale Means, Standard Deviations, and Coefficient Alphas

Sample Groups	Interest-Finder			Hypothetical Interest-Finder		
	Mean	SD	Alpha	Mean	SD	Alpha
Overall Sample	11.42	10.28	.95	9.93	10.82	.96
Female Sample	10.45	9.79	.95	8.51	10.27	.96
Black Sample	10.84	9.56	.94	9.96	10.24	.96
Hispanic Sample	10.38	9.63	.94	11.53	11.19	.96
Low SES Sample	10.31	9.57	.94	8.80	9.94	.96

The means for the Interest-Finder Investigative scale were all considerably larger than those for the Hypothetical Interest-Finder Investigative scale, ranging from about four-fifths of a point (Hispanic sample) to almost two points (Female sample). Unlike the Realistic scale comparison, there were no consistent subgroup differences between the two Investigative scales. While the differences between the Overall sample and the Low SES sample were about the same for both Investigative scales (1.11 and 1.13 points respectively), there were substantial Overall sample - Female sample differences (.97 and 1.42 points respectively) and Overall sample - Hispanic sample differences (1.04 and −1.60 points respectively) and substantial Overall sample - Black sample differences (.58 and −.03 points respectively). Thus, it appears that the Interest-Finder Investigative scale exhibited fewer subgroup differences for Females and Hispanics and larger subgroup differences for Blacks.

With regard to the standard deviations, it appears that the Hypothetical Interest-Finder Investigative scale had larger standard deviations for some subgroups (Females, Blacks, Hispanics) but about the same standard deviation for the Overall sample. The coefficient alphas for the Hypothetical Interest-Finder Investigative scale were only slightly larger than the coefficient alphas on the Interest-Finder Investigative scale--about .02 at most.

These findings indicate that the two Investigative scales are similar. This similarity can be seen in the high degree of item overlap (70%) and a similar coverage of the Investigative scale content categories. While there were some important subgroup differences, there was no consistent pattern. And although the Overall standard deviation was about the same for the two scales, the Hypothetical Interest-Finder Investigative scale tended to exhibit larger subgroup standard deviations.

Artistic Interest-Finder and Hypothetical Interest-Finder Scale Comparison

Cohen's kappa was computed to assess the degree to which the items selected for the Interest-Finder Artistic scale were also selected for inclusion in the Hypothetical Interest-Finder Artistic scale. While of moderate size ($k = .440$), kappa still suggested that the two forms shared significantly more items than would be expected by chance alone. The items that comprise both Artistic scales are reported in Table 10. Of the 40 items selected for each scale, there were 26 (65%) common items. Random selection would have led to 15 common items, so it appears that there was a high degree of item-overlap between the two Interest-Finder Artistic scales.

As was the case for the Investigative scale, the issues surrounding similarity of content coverage were still important for the Artistic scales. However, the relatively high degree of overlap between the two Artistic scales also suggested that this would not be a substantially large issue. As such, there was also less concern that the two scales might be assessing different aspects of the construct(s) that underlie the Artistic dimension or that the two Interest-Finder Artistic scales might exhibit different psychometric properties.

Table 10

Interest-Finder and Hypothetical Interest-Finder Artistic Scale Items

Item Type	Interest-Finder	Hypothetical Interest-Finder
Act	**Conduct a symphony orchestra**	**Conduct a symphony orchestra**
Act	**Design a set for a play**	**Design a set for a play**
Act	**Identify different styles of music(e.g., classical and rhythm and blues)**	**Identify different styles of music(e.g., classical and rhythm and blues)**
Act	**Join a local theater company**	**Join a local theater company**
Act	**Play the organ for a choir**	**Play the organ for a choir**
Act	**Write musical jingles for television commercials**	**Write musical jingles for television commercials**
Act	**Write script ideas for TV shows**	**Write script ideas for TV shows**
Act	**Write short stories**	**Write short stories**
Act	**Write the words for a song**	**Write the words for a song**
Act	Write a story about my favorite hobby	Arrange the subjects for a photograph
Act	Make figures out of clay	Sing in a musical
Act	Direct a play	Sing in a musical variety show
Act	Design toys for a manufacturer	Write a children's book
Act	Create pottery	Write a script for a play
Train	**Acting**	**Acting**
Train	**Creative writing**	**Creative writing**
Train	**Leading a band**	**Leading a band**
Train	**Music**	**Music**
Train	**Piano**	**Piano**
Train	**Reading music**	**Reading music**
Train	**Script writing**	**Script writing**
Train	Art history	Directing theatrical productions
Train	Ballet	Greeting card design
Train	Cartooning	Playing a musical instrument
Train	Making puppets	Singing
Train	Sculpting	Songwriting
Occ	**Author**	**Author**
Occ	**Guitarist**	**Guitarist**
Occ	**Illustrator**	**Illustrator**
Occ	**Jazz musician**	**Jazz musician**
Occ	**Magazine writer**	**Magazine writer**
Occ	**Movie critic**	**Movie critic**
Occ	**Photojournalist**	**Photojournalist**
Occ	**Portrait painter**	**Portrait painter**
Occ	**Puppeteer**	**Puppeteer**
Occ	**Songwriter**	**Songwriter**
Occ	Circus performer	Costume designer
Occ	Clown	Scriptwriter
Occ	Comedian	Set designer for a play
Occ	Rock band singer	Short story writer

Note. Items common to both scales are in boldface type. Act = Activities item. Train = Training item. Occ = Occupation item.

When the items that comprised the two Interest-Finder Artistic scales were categorized according to the Artistic content taxonomy (see Table 11), no significant differences emerged [$\chi^2(4, N = 80) = 4.95, p = .292$]. While there were some apparent differences in the coverage, particularly for Music and Visual Arts, it appeared that the content coverage was approximately the same for both Interest-Finder Artistic scales.

Table 11
Content Comparison of the Two Artistic Scales

	Interest-Finder Items		Hypothetical Interest-Finder Items	
Artistic Content Areas	Number	Percent	Number	Percent
Design and Decoration	2	5.0%	4	10.0%
Literary Works	8	20.0%	10	25.0%
Music	13	32.5%	18	45.0%
Performing Arts	8	20.0%	4	10.0%
Visual Arts	9	22.5%	4	10.0%

Table 12 reports the estimated means, standard deviations, and coefficient alphas for both of the Interest-Finder Artistic scales.

Table 12
Artistic Scale Means, Standard Deviations, and Coefficient Alphas

	Interest-Finder			Hypothetical Interest-Finder		
Sample Groups	Mean	SD	Alpha	Mean	SD	Alpha
Overall Sample	11.06	9.58	.94	11.25	10.58	.95
Female Sample	11.48	10.06	.95	12.54	11.12	.96
Black Sample	11.39	8.56	.92	12.20	9.60	.94
Hispanic Sample	11.92	8.77	.92	12.88	10.81	.95
Low SES Sample	10.48	9.02	.93	10.70	10.21	.95

Unlike the Realistic and Investigative areas, the means for the Interest-Finder Artistic scale were all smaller than those on the Hypothetical Interest-Finder Artistic scale, ranging from only about one-fifth of a point (Overall sample) to around one point (Female and Hispanic samples). Like the Realistic scale comparison, there were consistent subgroup differences (Females, Blacks, Hispanics) on the two Artistic scales. For each of these subgroups, the Hypothetical Interest-Finder Artistic scale exhibited larger subgroup differences than the Interest-Finder Artistic scale. While the absolute magnitude of the subgroup differences was relatively small for the Interest-Finder (M = .55), the difference was twice as large for the Hypothetical Interest-Finder (M = 1.11) In the same way, the Hypothetical Interest-Finder Artistic scale exhibited larger standard deviations than did the Interest-Finder Artistic scale. The differences in the standard deviations were generally about one point. Again, the coefficient alphas for Hypothetical Interest-Finder were only slightly larger than the coefficient alphas for their Interest-Finder counterparts--at most about .02.

The findings suggest that the Interest-Finder and Hypothetical Interest-Finder Artistic scales were similar. This similarity can be seen in the high degree of item overlap (65%), indicating similar coverage of the Artistic scale content categories. While there were some important subgroup differences, there was no consistent pattern in these differences. The Overall sample standard deviation was about the same for the two Artistic scales, even though the Hypothetical Interest-Finder scale tended to exhibit larger subgroup standard deviations than the Interest-Finder scale.

Social Interest-Finder and Hypothetical Interest-Finder Scale Comparison

Cohen's kappa was computed to assess the degree to which the items selected for the Interest-Finder Social scale were also selected for the Hypothetical Interest-Finder Social scale. The value of kappa was low ($k = .192$), which suggested that the two Social scales shared items at about the chance level. The items that comprised the two Social scales are reported in Table 13. Of the 40 items selected for each scale, there were only 21 (53%) common items. Random selection would have led to 17 common items. Consequently, it appears that there was little item-overlap between the two Interest-Finder Social scales. This again brings the issues of content coverage and potentially different psychometric properties to the forefront.

Table 13

Interest-Finder and Hypothetical Interest-Finder Social Scale Items

Item Type	Interest-Finder	Hypothetical Interest-Finder
Act	**Belong to a student organization**	**Belong to a student organization**
Act	**Conduct a training course**	**Conduct a training course**
Act	**Host social events**	**Host social events**
Act	**Lead a group therapy session**	**Lead a group therapy session**
Act	**Serve as a resident assistant in a college dormitory**	**Serve as a resident assistant in a college dormitory**
Act	**Teach adult education classes**	**Teach adult education classes**
Act	**Teach classes for new employees**	**Teach classes for new employees**
Act	**Tutor a student**	**Tutor a student**
Act	**Work as a camp counselor**	**Work as a camp counselor**
Act	**Work as a student tour guide for my school**	**Work as a student tour guide for my school**
Act	Assist a summer camp instructor	Administer first aid
Act	Volunteer for the Peace Corps	Counsel people having marital problems
Act	Work at a blood drive	Participate in Big Brother/Big Sister Program
Act	Work in a health club	Provide counseling to people with drug or alcohol problems
Train	**Braille (reading system for the blind)**	**Braille (reading system for the blind)**
Train	**Elementary school education**	**Elementary school education**
Train	**Interviewing and counseling methods**	**Interviewing and counseling methods**
Train	**Leading a support group**	**Leading a support group**
Train	Caring for the elderly	Caring for physically handicapped children
Train	First aid	Organizing recreational activities for teenagers
Train	Lifeguard training	Peer counseling
Train	Physical education	Teaching adults how to read
Train	Planning foreign travel	Teaching elementary school children
Train	Providing services to airplane passengers	Teaching learning-disabled students how to read
Train	Speaking a foreign language	Working as a camp counselor
Train	Teaching children how to swim	Working with abused children
Occ	**Adoption agency caseworker**	**Adoption agency caseworker**
Occ	**Camp counselor**	**Camp counselor**
Occ	**Drug and alcohol counselor**	**Drug and alcohol counselor**
Occ	**Interpreter for the hearing impaired**	**Interpreter for the hearing impaired**
Occ	**Special education teacher**	**Special education teacher**
Occ	**Teacher**	**Teacher**
Occ	**Weight-loss counselor**	**Weight-loss counselor**
Occ	Cruise activities director	Children's day-care worker
Occ	Dental hygienist	Day camp director
Occ	Flight attendant	Kindergarten teacher
Occ	High school principal	Nurse
Occ	Medical assistant	Planning foreign travel
Occ	Recreation coordinator	Red Cross volunteer
Occ	Sightseeing tour guide	Teacher aide

Note. Items common to both scales are in boldface type. Act = Activities item. Train = Training item. Occ = Occupation item.

When the items that comprised the two Social scales were categorized according to the Social scale taxonomy, no significant differences were found [c2(3, N = 80) = .33, p = .953]. So, even though there was little item overlap beyond chance expectations, both scales appeared to provide similar coverage of the Social content, as reported in Table 14.

Table 14
Content Comparison of the Two Social Scales

Social Content Areas	Interest-Finder Items		Hypothetical Interest-Finder Items	
	Number	Percent	Number	Percent
Organizing Group Activities	5	12.5%	6	15.0%
Personal Services	11	27.5%	9	22.5%
Provision of Human Services	11	27.5%	12	30.0%
Teaching and Education	13	32.5%	13	32.5%

Table 15 reports the estimated means, standard deviations, and coefficient alphas (Lord & Novick, 1968) for both Social scales.

Table 15
Social Scale Means, Standard Deviations, and Coefficient Alphas

Sample Groups	Interest-Finder			Hypothetical Interest-Finder		
	Mean	SD	Alpha	Mean	SD	Alpha
Overall Sample	13.18	9.59	.93	14.28	11.48	.96
Female Sample	15.72	9.24	.92	18.45	10.83	.94
Black Sample	14.17	9.20	.92	15.90	10.83	.95
Hispanic Sample	13.16	9.88	.94	13.88	11.44	.96
Low SES Sample	12.76	9.39	.93	14.17	11.36	.96

The means for the Interest-Finder Social scale were all smaller than those for the Hypothetical Interest-Finder Social scale, with the differences ranging from only about seven-tenths of a point (Hispanic sample) to around two and a half points (Female sample). Given the Overall sample difference between the two Social scales (1.10 points), it is not surprising that there were large subgroup differences for the Female (2.73 points), Black (1.73 points) and Low SES (1.41 points) samples. Perhaps more important, it appears that there were pattern differences as well. On the Interest-Finder Social scale, the Female sample mean was well above the Overall sample mean (2.54 points); this difference was much larger for the Hypothetical Interest-Finder Social scale (4.17 points). While not as large, this same finding was observed for the Black sample, which scored a point (.99) higher than the Overall sample on the Interest-Finder Social scale, and closer to two points higher (1.62) on the Hypothetical Interest-Finder Social scale.

A comparison of the standard deviations leads to similar conclusions, with Hypothetical Interest-Finder scale values uniformly larger than their respective Interest-Finder scale values. This difference averaged close to a point and a half (1.44). Again, the alphas were uniformly and slightly higher for the Hypothetical Interest-Finder Social scale than for the Interest-Finder Social scale, with .03 as the largest difference.

Enterprising Interest-Finder and Hypothetical Interest-Finder Scale Comparison

Cohen's kappa was computed to assess the degree to which the items selected for the Interest-Finder Enterprising scale were also selected for the Hypothetical Interest-Finder Enterprising scale. The value of kappa was moderately low ($k = .309$), which indicated that the two scales shared items at about the chance level. The items that comprised both forms are reported in Table 16. Of the 40 items selected for each scale, there were only 24 (60%) common items. Random selection would have led to 17 common items, so it appeared that there was little item-overlap between the two Interest-Finder Enterprising scales. Not surprisingly, this again brings the issues of content coverage and potentially different psychometric properties to the forefront.

Table 16

Interest-Finder and Hypothetical Interest-Finder Enterprising Scale Items

Item Type	Interest-Finder	Hypothetical Interest-Finder
Act	**Develop a plan to boost the sales of a product**	**Develop a plan to boost the sales of a product**
Act	**Manage a department within a company**	**Manage a department within a company**
Act	**Manage the sales of a large corporation**	**Manage the sales of a large corporation**
Act	**Persuade management to see the employees' side of a debate**	**Persuade management to see the employees' side of a debate**
Act	**Present a new advertising campaign to corporate executive**	**Present a new advertising campaign to corporate executive**
Act	**Sell plans to develop new areas of real estate**	**Sell plans to develop new areas of real estate**
Act	**Think of an idea to start a new business**	**Think of an idea to start a new business**
Act	**Write up contracts between two parties**	**Write up contracts between two parties**
Act	Argue in favor of a new law	Convince people to vote for a candidate
Act	Convince others that my ideas/suggestions are best	Lead a seminar on taking business risks
Act	Convince people to follow my lead	Manage a new area of a large corporation
Act	Debate with others about politics	Run for public office
Act	Manage a restaurant	Think of ideas for starting your own business
Act	Present a case in front of a judge and jury	Work as a salesperson in a store
Train	**Developing leadership skills**	**Developing leadership skills**
Train	**Effective marketing strategies**	**Effective marketing strategies**
Train	**Identifying new business opportunities**	**Identifying new business opportunities**
Train	**Organizing the work of several people**	**Organizing the work of several people**
Train	**Project management**	**Project management**
Train	**Starting your own business**	**Starting your own business**
Train	Buying and selling stock	Applying for a small business loan
Train	Developing business plans	Changing the structure of a corporation
Train	How to succeed in the corporate world	Developing effective presentations
Train	Importing and exporting goods for a profit	Finding people to invest in your business
Train	Law	Helping your business grow
Train	Managing an organization or business	Managing a political campaign
Occ	**Agent for actors**	**Agent for actors**
Occ	**Auctioneer**	**Auctioneer**
Occ	**Company spokesperson**	**Company spokesperson**
Occ	**Corporate executive**	**Corporate executive**
Occ	**District attorney**	**District attorney**
Occ	**Hotel manager**	**Hotel manager**
Occ	**Mayor**	**Mayor**
Occ	**Retail store owner**	**Retail store owner**
Occ	**Supreme Court justice**	**Supreme Court justice**
Occ	**Traveling salesperson**	**Traveling salesperson**
Occ	Cosmetics sales representative	Apartment complex manager
Occ	Gift shop owner	Health club director
Occ	Lawyer for sports figures and movie stars	Politician
Occ	Owner of a professional sports team	Real estate developer

Note. Items common to both scales are in boldface type. Act = Activities item. Train = Training item. Occ = Occupation item.

When the items that comprised the two scales were categorized according to the Enterprising scale taxonomy, significant differences failed to emerge [$\chi^2(4, N = 80) = 2.03, p = .731$]. So, even though there was little item overlap beyond chance expectations, both Enterprising scales seemed to provide similar coverage of the Enterprising content, as reported in Table 17.

Table 17
Content Comparison of the Two Enterprising Scales

	Interest-Finder Items		Hypothetical Interest-Finder Items	
Enterprising Content Areas	Number	Percent	Number	Percent
Business Venture	10	25.0%	9	22.5%
Buying, Selling and Persuasion	12	30.0%	11	27.5%
Law	6	15.0%	3	7.5%
Leadership and Management	9	22.5%	12	30.0%
Public Speaking/Public Relations and Politics	3	7.5%	5	12.5%

Table 18 reports the estimated means, standard deviations, and coefficient alphas for both Enterprising scales.

Table 18
Enterprising Scale Means, Standard Deviations, and Coefficient Alphas

	Interest-Finder			Hypothetical Interest-Finder		
Sample Groups	Mean	SD	Alpha	Mean	SD	Alpha
Overall Sample	12.80	10.68	.95	11.44	10.96	.96
Female Sample	12.99	10.62	.95	11.44	10.80	.96
Black Sample	14.35	10.87	.95	13.36	11.22	.96
Hispanic Sample	12.60	9.59	.94	12.12	10.49	.95
Low SES Sample	12.20	10.81	.95	10.78	11.09	.96

The means for the Interest-Finder Enterprising scale appeared to be larger than their Hypothetical Interest-Finder Enterprising scale counterparts. These mean differences ranged from a low of half-a-point (.48) for the Hispanic sample to a high of about a point-and-a-half (1.55) for the Female sample, with the Overall sample differences being in between (1.36 points). For both scales, there was a large Overall sample-Black sample difference (−1.55 and −1.92 points respectively). The two Interest-Finder Enterprising scales appeared to have quite similar standard deviations and coefficient alphas.

The item overlap, taxonomic coverage, and psychometric considerations suggested few real differences between these two Interest-Finder Enterprising scales beyond the tendency for the Interest-Finder Enterprising scale to exhibit relatively larger subgroup scale means than those of their respective Hypothetical Interest-Finder Enterprising scale subgroups.

Conventional Interest-Finder and Hypothetical Interest-Finder Scale Comparison

Cohen's kappa was computed to assess the degree to which the items selected for the Interest-Finder Conventional scale were also selected for inclusion in the Hypothetical Interest-Finder Conventional scale. The value of kappa was moderate (k = .400), which indicated that the two scales shared items only at about the chance level. The items that comprised both forms are reported in Table 19. Of the 40 items selected for each scale, there were only 26 (65%) common items. Random selection would have led to 17 common items, suggesting that there was relatively little item overlap between the two Interest-Finder Conventional scales beyond that attributable to chance expectations. As was the case with some previous areas, this brings the issues of content coverage and potentially different psychometric properties to the forefront.

Table 19

Interest-Finder and Hypothetical Interest-Finder Conventional Scale Items

Item Type	Interest-Finder	Hypothetical Interest-Finder
Act	**Complete forms for items to be shipped**	**Complete forms for items to be shipped**
Act	**Count the inventory of a small business**	**Count the inventory of a small business**
Act	**Enter data into a computer**	**Enter data into a computer**
Act	**Keep accurate financial records for an organization**	**Keep accurate financial records for an organization**
Act	**Make entries into a financial accounting system**	**Make entries into a financial accounting system**
Act	**Organize and maintain files**	**Organize and maintain files**
Act	**Put accurate price tags on merchandise**	**Put accurate price tags on merchandise**
Act	**Review financial records of an organization**	**Review financial records of an organization**
Act	**Set up and maintain a filing system**	**Set up and maintain a filing system**
Act	Learn the major sections of a business letter	Filling out insurance claim forms
Act	Operate a telephone switchboard	Help new employees fill out insurance forms
Act	Sort mail	Improve a small business accounting system
Act	Type reports	Keep records of goods sold each day at a store
Act	Weigh packages to determine postage due	Sort and alphabetize files
Occ	**Accountant**	**Accountant**
Occ	**Accounting clerk**	**Accounting clerk**
Occ	**Court clerk**	**Court clerk**
Occ	**Data entry clerk**	**Data entry clerk**
Occ	**Mail room clerk**	**Mail room clerk**
Occ	**Payroll specialist**	**Payroll specialist**
Occ	**Reservation clerk**	**Reservation clerk**
Occ	**Scheduler**	**Scheduler**
Occ	**Supply and inventory specialist**	**Supply and inventory specialist**
Occ	**Tax preparer**	**Tax preparer**
Occ	Computer operator	Auditor
Occ	Office assistant	Billing clerk
Occ	Personnel clerk	Supply room clerk
Occ	Word processing operator	Tax accountant
Train	**Determining yearly taxes for companies**	**Determining yearly taxes for companies**
Train	**Personnel records management**	**Personnel records management**
Train	**Preparing budgets**	**Preparing budgets**
Train	**Basic accounting principles**	**Basic accounting principles**
Train	**Formatting a letter correctly**	**Formatting a letter correctly**
Train	**Recordkeeping systems**	**Recordkeeping systems**
Train	**Stock control and accounting procedures**	**Stock control and accounting procedures**
Train	Balancing a checkbook	Computing wages for payroll records
Train	Maintaining a computer data base	Entering data into a computer
Train	Operating photocopying machines	Office filing system design
Train	Procedures for the handling and storage of goods	Preparing tax withholding forms for new employees
Train	Warehouse inventory systems	Tax accounting

Note. Items common to both scales are in boldface type. Act = Activities item. Train = Training item. Occ = Occupation item.

When the items that comprised the two Conventional scales were categorized according to the Conventional scale taxonomy, no significant differences were found [$\chi^2(3, N = 80) = 4.62, p = .202$]. So, even though there was little item overlap beyond chance expectations, both forms seemed to provide similar coverage of the Conventional content, as reported in Table 20.

Table 20
Content Comparison of the Two Conventional Scales

Conventional Content Areas	Interest-Finder Items		Hypothetical Interest-Finder Items	
	Number	Percent	Number	Percent
Bookkeeping and accounting	11	27.5%	15	37.5%
Filing, Record Keeping and Inventory Control	16	40.0%	20	50.0%
Operate Office Machinery	8	20.0%	3	7.5%
Typing, Word Processing and General Clerical	5	12.5%	2	5.0%

Table 21 reports the estimated means, standard deviations, and coefficient alphas for both of the Interest-Finder Conventional scales. It appeared that the means for the Interest-Finder Conventional scale were considerably larger than the means for their respective subgroups on the Hypothetical Interest-Finder Conventional scale. For both scales, there was a large Overall sample-Black sample difference (–2.44 and –2.74 points respectively). The two forms appeared to have very similar standard deviations and coefficient alphas.

Table 21
Conventional Scale Means, Standard Deviations, and Coefficient Alphas

Sample Groups	Interest-Finder			Hypothetical Interest-Finder		
	Mean	SD	Alpha	Mean	SD	Alpha
Overall Sample	9.67	10.56	.96	8.92	10.88	.97
Female Sample	10.88	11.25	.97	10.22	11.68	.97
Black Sample	12.11	11.00	.96	11.66	11.09	.96
Hispanic Sample	9.84	9.60	.95	9.48	10.42	.96
Low SES Sample	10.20	10.54	.96	9.44	10.92	.97

The item overlap, taxonomic coverage, and psychometric considerations suggested few real differences between these two Interest-Finder Conventional scales beyond the tendency for the Interest-Finder Conventional scale to have relatively larger scale means than the Hypothetical Interest-Finder Conventional scale.

Utilization of Results

Conclusions Based on Scale-Level Comparisons

In this section, conclusions are offered as to which version of the Interest-Finder is a better version, Interest-Finder (with pattern matching) or Hypothetical Interest-Finder (without pattern matching). For each of the RIASEC domains, taxonomic and psychometric comparisons were made between the two versions of the Interest-Finder. Table 22 summarizes these comparisons, and shows the areas for which one or the other version was superior.

Table 22
Summary of the Findings of the Scale Comparisons: Which is the Better Version of the Interest-Finder?

	Type of Consideration				
RIASEC Scales	Percent Item Overlap	Cohen's Kappa	Content Coverage	Psychometric Alpha	Psychometric Subgroup Differences
Realistic Scale	40%	.025	**I**[b]	*n*	**I**
Investigative Scale	70%	.524[a]	*n*	*n*	**I**
Artistic Scale	65%	.440[a]	**I**	*n*	**I**
Social Scale	53%	.192	*n*	*n*	**I**
Enterprising Scale	60%	.309	*n*	*n*	*n*
Conventional Scale	65%	.400	**I**	*n*	*n*

Note. For each comparison, I indicates that the Interest Finder was the better scale and *n* indicates that neither scale was the better scale. [a] Kappa was statistically significant beyond the .05 level, indicating that the item overlap was beyond what would be expected based on random item selection. [b] Interest-Finder provided a significantly better coverage of the taxonomy than did the Hypothetical Interest-Finder.

This summary comparison makes clear that the two Interest-Finder forms shared a number of common items, though the percentage of overlap was statistically significant for only the Investigative and Artistic areas. It is equally clear that the Interest-Finder tended to provide a more complete coverage of the RIASEC domains than did the Hypothetical Interest-Finder. This was true for the Realistic, Artistic, and Conventional domains. For the other three scales, neither version provided more complete coverage. Even so, this means that the Interest-Finder provided superior content coverage for half of the areas assessed by the measure. While the coefficient alphas for the Hypothetical Interest-Finder scales were uniformly larger than the coefficient alphas for the Interest-Finder scales, these differences were quite negligible--never exceeding .04 in magnitude. Since all of the scales had alphas of at least .91, there is no clear reason to prefer one version over the other. On the other hand, there is a clear reason to prefer Interest-Finder over Hypothetical Interest-Finder when the criterion is that of subgroup differences. In those RIASEC scales for which there were clear differences in the magnitude of the subgroup mean differences between the two versions, Interest-Finder manifested the smaller differences. This was especially true in the comparisons involving the Female sample with the Overall sample.

With these considerations in mind, it is reasonable to conclude that the scale-level comparisons endorse Interest-Finder as the better of the two versions.

Conclusions Based on Form Comparisons

Now that the two Interest-Finder versions have been compared in a scale-by-scale fashion, it is important to compare the two versions at the form level. Overall, the two versions shared 141 (58.5%) common items, a number that is half-again-larger than what would be expected by chance alone (95 items). Even so, kappa was relatively low ($k = .318$), which suggests that while the overlap was larger than what might be expected due to random selection of the items for the two versions, the two versions did not share common items significantly beyond what would be expected based on random item selection. From this it might be surmised that the two Interest-Finder versions are, indeed, fundamentally different instruments. In this regard it is reasonably safe to conclude that the SDS pattern matching did have an important influence on the items that were selected for inclusion in the Interest-Finder.

More importantly, however, is whether such pattern matching adversely affected the coverage of the relevant content areas in the RIASEC constructs. A statistical comparison of the two Interest-Finder forms is not feasible, due to the violation of the minimum expected number of observations per cell. However, an examination of the results of the scale-by-scale comparisons do allow conclusions about the influence of the SDS pattern matching on content considerations. There were no differences in the content coverage on five of the six RIASEC scales. The one difference was on the Realistic scale, for which pattern matching seemed to lead to a more thorough coverage of the content. Based on content considerations, it is reasonable to conclude that pattern matching had a positive, rather than negative, influence on the item-selection process utilized to select items for inclusion in the Interest-Finder.

An analysis of the expected means, standard deviations, and coefficient alphas which were calculated from the item-tryout data for both forms of the Interest-Finder showed that, again, pattern matching either negligibly influenced the results or positively influenced the results.

Together, these findings argue quite substantially that SDS pattern matching led to the creation of a better version of the Interest-Finder than would have been created if pattern matching had not been effected. Based on these considerations, it appears that the influence of the SDS pattern matching on the Interest-Finder was not negative and that, at worst, the current version of the Interest-Finder suffers no ill effects because of that pattern matching.

REFERENCES

Anastasi, A. (1988). *Psychological testing* (6th ed.). New York: Macmillan Publishing.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cronbach, L. J. (1990). *Essentials of psychological testing* (5th ed.).New York: Harper and Row.

Eagle, E. (1989). *Socioeconomic status, family structure, and parental involvement: The correlates of achievement.* ERIC ED307332.

Hansen, J. C., & Campbell, D. P. (1985). *Manual for the Strong Interest Inventory* (4th ed.) Palo Alto, CA: Consulting Psychologists Press, Inc.

Holland, J. L. (1973). *Making vocational choices: A theory of vocational personalities and work environments.* Englewood Cliffs, NJ: Prentice-Hall.

Holland, J. L. (1985a). *The Self-Directed Search professional manual.* Odessa, FL: Psychological Assessment Resources.

Holland, J. L. (1985b). *Making vocational choices: A theory of vocational personalities and work environments* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores.* Reading, MA: Addison-Wesley.

Defense Manpower Data Center. (1994). *Technical manual for the ASVAB 18/19 Career Exploration Program.* Seaside, CA: Author.

Wall, J. E. (1994). An example of assessment's role in career exploration. *Journal of Counseling and Development, 72*, 608-613.

Wall, J. E., Wise, L. L., & Baker, H. E. (in press). Development of the Interest-Finder - a new RIASEC-based interest inventory. *Measurement and Evaluation in Counseling and Development.*

Wall, J. E., Wise, L. L., & Baker, H. E. [*Interest-Finder Differences Due to Grade-Level*] Unpublished raw data.

APPENDIX A

INTEREST-FINDER FORM ASSEMBLY PROGRAM NOTES

Interest-Finder Form Assembly Program Notes

(Notes by I. A. Krass)

Let us have a set Y of m items, where each item $y_i \in Y; i = 1,...,m$ is a vector of the item property. Currently, there are 11 dimensions, each representing one unique item property, such that $y_i = (y_1^i,... y_{11}^i)$, where:

y_1^i - Total sample endorsement rate;

y_2^i - Female sample endorsement rate;

y_3^i - Black sample endorsement rate;

y_4^i - Hispanic sample endorsement rate;

y_5^i - Low SES sample endorsement rate;

y_6^i -Item-biserial correlation with SDS Realistic scale;

y_7^i -Item-biserial correlation with SDS Investigative scale;

y_8^i - Item-biserial correlation with SDS Artistic scale;

y_9^i - Item-biserial correlation with SDS Social scale;

y_{10}^i - Item-biserial correlation with SDS Enterprising scale;

y_{11}^i - Item-biserial correlation with SDS Conventional.

All items are divided into six disjoint subsets according to the RIASEC taxonomy $Y_j; j = 1,...,6$, such that $Y = \bigcup Y_j$, and $Y_{j_1} \neq Y_{j_2}$, if $j_1 \neq j_2$. Here:

Y_1 - set of all Interest-Finder Realistic items;

Y_2 - set of all Interest-Finder Investigative items;

Y_3 - set of all Interest-Finder Artistic items;

Y_4 - set of all Interest-Finder Social items;

Y_5 - set of all Interest-Finder Enterprising items;

Y_6 - set of all Interest-Finder Conventional items.

Moreover, every RIASEC taxonomy set is further subdivided into three disjoint taxonomy subsets. These are characterized by triple indices (l, p, q), where l corresponds to Activity items, p to Training items, and q to Occupation items. Thus,

$$Y_j = \bigcup Y_{lpq}^j,\ j = 1,...,6,\ \text{and}\ Y_{l_1 p_1 q_1}^j \neq Y_{l_2 p_2 q_2}^j \ \text{if}\ (l_1, p_1, q_1) \neq (l_2, p_2, q_2).$$

The Form Assembly program consists of creating a testing subset

$\overline{\mathrm{Y}} = \{y_i, i = 1,\ldots,n\}$ from the original item pool; i. e. $\overline{\mathrm{Y}} \subset \mathrm{Y} = \{y_i, i = 1,\ldots,m\}$,

such that the items in a testing subset should satisfy the taxonomy constraints and maximize an objective function to be defined shortly. The creation of set $\overline{\mathrm{Y}}$ is accomplished through an assignment function $x_i \in \{0,1\}; i = 1,\ldots,m.$, such that if an item i is assigned to the testing subset $\overline{\mathrm{Y}}$, then $x_i = 1$, and if item i is not so assigned, then $x_i = 0$. Obviously, the assignment function x_i should satisfy the equality:

$$\sum_{i=1}^{m} x_i = n \qquad (1)$$

which means that the testing subset $\overline{\mathrm{Y}}$ is filled. The assignment variables x_i should also satisfy all of the taxonomy constraints. First of all, they should satisfy the RIASEC taxonomy constraint:

$$\sum_{i \in \mathrm{Y}_j} x_i = \mathrm{A};\ j = 1,\ldots,6. \qquad (2)$$

(Currently $\mathrm{A} = 40.$) This constraint requires equal representative of all type of RIASEC items in the feasible assignment.

Together with (2), the assignment function x_i should satisfy these taxonomy constraints:

$$\sum_{p,q} \sum_{i \in \mathrm{Y}^i_{lpq}} x_i = \mathrm{B}_1;\ j = 1,\ldots,6, \qquad (3)$$

$$\sum_{l,q} \sum_{i \in \mathrm{Y}^i_{lpq}} x_i = \mathrm{B}_2;\ j = 1,\ldots,6, \qquad (4)$$

$$\sum_{l,p} \sum_{i \in \mathrm{Y}^i_{lpq}} x_i = \mathrm{B}_3;\ j = 1,\ldots,6. \qquad (5)$$

Equality (3) corresponds to the Activity items taxonomy constraint (currently $B_1 = 14$); equality (4) corresponds to Training items taxonomy constraint (currently $B_2 = 12$); and equality (5) corresponds to Occupation items taxonomy constraint (currently $B_3 = 14$).

ut of all of the feasible assignments satisfying (1) - (5) x_i we should choose an optimal assignment $\bar{x}_i$ which maximizes the objective function:

$$\Phi(\{\bar{x}_i\}) = \sum_{k=1}^{6} w_k \cdot \Phi_k(\{\bar{x}_i\}). \qquad (6)$$

Here $\Phi_k(\{x_i\}),\ k = 1,\ldots,6$ are the subobjective functions, and $w_k \in (0,1)$ are the chosen weights of those subobjective functions in the global optimization. These subobjective functions are:

$$\Phi_1(\{x_i\}) = \sum_{i=1}^{m} x_i \cdot y_1^i$$

used to maximize the item-target SCS scale correlation;

$$\Phi_2(\{x_i\}) = -\sum_{i=1}^{m} x_i \cdot (y_2^i - y_3^i)$$

used to minimize the difference between the Overall sample endorsement rate and the Female sample endorsement rate;

$$\Phi_3(\{x_i\}) = -\sum_{i=1}^{m} x_i \cdot (y_2^i - y_4^i)$$

used to minimize the difference between the Overall sample endorsement rate and the Black sample endorsement rate;

$$\Phi_4(\{x_i\}) = -\sum_{i=1}^{m} x_i \cdot (y_2^i - y_5^i)$$

used to minimize the difference between the Overall sample endorsement rate and Hispanic sample endorsement rate;

$$\Phi_5(\{x_i\}) = -\sum_{i=1}^{m} x_i \cdot (y_2^i - y_6^i)$$

used to minimize the difference between the Overall sample endorsement rate and the Low SES sample endorsement rate.

The program (1) - (6) can be converted to a Linear Programming problem and solved as a mixed integer problem by a corresponding commercial optimizer, or a heuristic solution can be found, for example, by a version utilizing a greedy algorithm such as was used in this situation.